

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 02:50:58 ; Search time 351 Seconds

(without alignments)
6935.642 Million cell updates/sec

Title: US-09-787-844-1

Perfect score: 1081
Sequence: 1 agagagcagcagggggggcgtca.....acattccagttgatgctctgc 1081

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

N.Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1081	100.0	1081	20	AA87151
2	1081	100.0	1081	20	AA086505
3	1063.8	98.9	1087	21	AA62015
4	1063.2	98.4	1131	22	AA501698
5	1055.8	97.7	1100	20	AA52259
6	1055.8	97.7	1100	22	AA72417
7	1054.8	97.6	1077	24	AA27418
8	1054.8	97.4	1079	21	AA45673
9	1052.4				

10	1049.6	97.1	1100	19	AAV59119	Nucleotide sequenc
11	1048.6	97.0	1117	19	AAV59118	Nucleotide sequenc
12	1045.8	96.7	1082	20	AAV15337	Nucleic acid encod
13	1045.8	96.7	1082	24	AA27417	Human eosinophil s
14	1037.8	96.0	1082	24	AA502648	Human secreted pro
15	938.8	86.8	942	20	AAV15336	Nucleic acid encod
16	938.8	86.8	942	21	AA62005	Nucleotide sequenc
17	752.6	69.6	1103	22	AAV55271	Hydrophobic domain
18	496.4	45.9	498	22	AA501708	WISH TADG-16 catal
19	494.8	45.8	498	22	AA501707	WISH TADG-16 catal
20	493.4	45.6	529	24	AA561822	Lung small cell ca
21	479	44.3	959	19	AAV59132	Nucleotide sequenc
22	475	43.9	918	22	AA167198	Nucleotide sequenc
23	465.4	43.1	575	22	AA524506	Human ovarian tumo
24	465.4	43.1	575	22	AAH83114	Human ovarian tumo
25	422.2	39.1	888	24	ABK31769	DNA encoding novel
26	412.2	38.2	1018	24	AA27415	Human eosinophil s
27	412.2	38.2	1018	24	AAV59133	Genomic sequence o
28	329.8	30.5	837	24	AA27419	Human eosinophil s
29	318.2	29.4	373	24	AB180716	Human ovarian can
30	165	15.3	1110	22	AA766994	Human protease T c
31	165	15.3	1129	22	AA521354	Human cDNA sequenc
32	165	15.3	1151	22	AA208286	Human secreted pro
33	165	15.3	1703	21	AA52473	Human cDNA sequenc
34	163.2	15.1	980	19	AAV59136	Nucleotide sequenc
35	161.8	15.0	1212	22	AA158844	Human polynucleoti
36	161.8	14.9	2457	24	ABK31795	DNA encoding novel
37	157.2	14.5	1157	22	AA160330	Human polynucleoti
38	155.8	14.4	1150	22	AAV77000	Fusion gene of pro
39	151.8	14.0	1322	21	AA61704	CDNA encoding mous
40	150.6	13.9	2662	24	AA23854	Human protease PRT
41	150.6	13.9	2810	22	AAK94500	Human full-length
42	149	13.8	1958	24	ABK12891	Human secreted PRT
43	145.2	13.4	1122	24	AA27036	Human tryptase CDN
44	145	13.4	437	24	AB181643	Human eosinophil s
45	133.4	12.3	393	24	AA27416	Human eosinophil s

ALIGNMENTS

RESULT 1	AA87151	standard; cDNA: 1081 BP.
ID	AA87151	
AC	AA87151	
XX	27-SEP-1999	(first entry)
DT	27-SEP-1999	
DE	Human protease HUPM-3 CDNA.	
XX		
XX	Protease; human; HUPM-3; cell proliferation; cancer;	
KW	Immune disorder; inflammation; therapy; ss.	
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
FT	CDS	40..984
FT		/*tag- a
FT	sig_peptide	40..76
FT		/*tag- b
FT	/note= "putative signal peptide sequence"	
FT	mat_peptide	77..981
FT		/*tag- b
PN	MO9936550-A2.	
XX		
XX	22-JUL-1999.	
PD		
XX		
XX	12-JAN-1999;	99WO-US00655.
PF		
XX		
XX	Human PRO303 CDNA.	
PR	16-JAN-1998;	98US-0008271.
XX		

Exhibit A

Page 1 of 2

FT		/note= "protein kinase C phosphorylation site"
FT	Modified-site	73
FT	Active-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	82
FT	Modified-site	86
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	127
FT	Modified-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	134
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	161
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	167
FT	Modified-site	/note= "N-glycosylated"
FT	Modified-site	190
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	200
FT	Modified-site	/note= "N-glycosylated"
FT	Modified-site	212
FT	Modified-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	238
FT	Modified-site	273
FT	Modified-site	/note= "N-glycosylated"
FT	Modified-site	291
FT	Modified-site	/note= "casein kinase II phosphorylation site"
XX		
PX	W09936550-A2.	
PD	22-JUN-1999.	
PF	12-JAN-1999;	99WC-US000655.
PR	16-JAN-1998;	98US-00008271.
PA	(INCY-) INCYTE PHARM INC.	
PZ	Bandman O, Hillman JL, Yue H, Guegler KJ, Corley NC, Tang YT,	
P1	Shan P;	
DR	WPI; 1999-430616/36.	
DR	N-PESDB; AAX87151.	
PT	Novel human protease molecules useful in the treatment of developmental disorders and/or cancers.	
PS	Claim 1; Page 71-72; 90PP; English.	
XX		
CC	The present sequence represents novel human protease HUMPM-3, as deduced from the consensus sequence (see AAX87151) of overlapping cDNA clones obtained from various libraries. Northern analysis shows expression of HUMPM-3 in cardiovascular, hematopoietic and male reproductive cDNA libraries. Approximately 86% of these libraries are associated with neoplastic disorders. The invention provides 12 new human proteases, i.e. HUMPM-1 to -12 (see AAY06432-43), and the polynucleotides encoding them (see AAX87149-60). Also provided are vectors, host cells and methods for producing HUMPM polypeptides, as well as agonists and antagonists of HUMPM. Methods for treating or preventing cell proliferative disorders and immune disorders using HUMPM or HUMPM antagonists are claimed	
SQ	Sequence 314 AA;	
	Query Match	96.4%; Score 1728; DB 2; Length 314;
	Best Local Similarity	100.0%; Pred. No. 5,8e-148;
	Matches 314; Conservative	0; Mismatches 0; Indels 0; Gaps 0
DY	14 MGARGALLIALIADAGLRKPESSQAAPLSGCGRRVITSRIVGEBDAELGWPMWGSLR	73
DB	1 MGARFALLIALILALAAAGLRKPESCEAAPLSGPCGGRVITSRIVGEDAELGWPMWGSLR	60
DY	74 LMDSHVCAGVSLISHRWALTAAHCFETYSGLSPSCGMVOFGQLTSMPSFWSLOAVYTRYE	133
DB	61 LMDSHVCCGVSLISHRWALTAAHCFETYSGLSPSCGMVOFGQLTSVPFSWLOAVYTRIF	120

Exhibit B

Page 2 of 2

QY 134 VSNITLSPRYLGNSTPYDIALVKLSAPVTYTHIQIPICQASTFEFENRTDCWVTGWSGK 193
 DB 121 VSNITLSPRYLGNSTPYDIALVKLSAPVTYTHIQIPICQASTFEFENRTDCWVTGWSGK 180
 QY 194 EDEALPSPHTLQEVQVAILINNSMCNHLFLKXSPFRDIFGDMVACGNNQGGDAFCGDSGG 253
 DB 181 EDEALPSPHTLQEVQVAILINNSMCNHLFLKXSPFRDIFGDMVACGNNQGGDAFCGDSGG 240
 QY 254 PLACNKGKLMWQIGVSWGCGRRPRPGVYTNISHEFEMIQKMAQSGMSQPPSPMLL 313
 DB 241 PLACNKGKLMWQIGVSWGCGRRPRPGVYTNISHEFEMIQKMAQSGMSQPPSPMLL 300
 QY 314 FFFLLMALPLLGPV 327
 DB 301 FFFLLMALPLLGPV 314
 RESULT 5
 AAY13388
 ID AAY13388 standard; protein; 314 AA.
 AC AAY13388;
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of protein PRO303.
 XX
 KM Secreted protein; transmembrane protein; human; enterocolitis;
 KM Zollinger-Ellison syndrome; gastrointestinal ulceration;
 KM congenital microvillus atrophy; skin disease; cell growth;
 KM abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
 KM Parkinson's disease; Alzheimer's disease; ALS; neuropathy; fibromodulin;
 KM dermal scarring; Usher Syndrome; Atrophla areata; and thrombotic;
 KM wound healing; tissue repair.
 XX
 XX Homo sapiens.
 XX OS
 XX MO9914328-A2.
 XX
 PD 25-MAR-1999.
 XX
 PF 16-SEP-1998; 98WO-US019330.
 XX
 PR 17-SEP-1997; 97US-0059113P.
 PR 17-SEP-1997; 97US-0059115P.
 PR 17-SEP-1997; 97US-0059117P.
 PR 17-SEP-1997; 97US-0059119P.
 PR 17-SEP-1997; 97US-0059121P.
 PR 17-SEP-1997; 97US-0059123P.
 PR 18-SEP-1997; 97US-0059125P.
 PR 18-SEP-1997; 97US-0059265P.
 PR 15-OCT-1997; 97US-0062125P.
 PR 17-OCT-1997; 97US-0062285P.
 PR 17-OCT-1997; 97US-0062287P.
 PR 21-OCT-1997; 97US-0063466P.
 PR 24-OCT-1997; 97US-0062814P.
 PR 24-OCT-1997; 97US-0062816P.
 PR 24-OCT-1997; 97US-0063045P.
 PR 24-OCT-1997; 97US-0063120P.
 PR 24-OCT-1997; 97US-0063121P.
 PR 24-OCT-1997; 97US-0063127P.
 PR 24-OCT-1997; 97US-0063128P.
 PR 27-OCT-1997; 97US-0063327P.
 PR 27-OCT-1997; 97US-0063329P.
 PR 28-OCT-1997; 97US-0063541P.
 PR 28-OCT-1997; 97US-0063542P.
 PR 28-OCT-1997; 97US-0063544P.
 PR 28-OCT-1997; 97US-0063549P.
 PR 28-OCT-1997; 97US-0063550P.
 PR 28-OCT-1997; 97US-0063556P.
 PR 29-OCT-1997; 97US-0063435P.
 PR 29-OCT-1997; 97US-0063704P.

PR 29-OCT-1997; 97US-0063732P.
 PR 29-OCT-1997; 97US-0063734P.
 PR 29-OCT-1997; 97US-0063735P.
 PR 29-OCT-1997; 97US-0063738P.
 PR 29-OCT-1997; 97US-0064215P.
 PR 31-OCT-1997; 97US-0063870P.
 PR 31-OCT-1997; 97US-0064103P.
 PR 03-NOV-1997; 97US-0064248P.
 PR 07-NOV-1997; 97US-0064809P.
 PR 12-NOV-1997; 97US-0065186P.
 PR 17-NOV-1997; 97US-0065845P.
 PR 18-NOV-1997; 97US-0065925P.
 PR 21-NOV-1997; 97US-0066120P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 24-NOV-1997; 97US-0066453P.
 PR 24-NOV-1997; 97US-0066466P.
 PR 24-NOV-1997; 97US-0066511P.
 PR 24-NOV-1997; 97US-0066770P.
 PR 24-NOV-1997; 97US-0066772P.
 PR 25-NOV-1997; 97US-0066840P.
 PA (GETH) GENENTECH INC.
 XX
 PI Wood WI, Gurney AL, Goddard A, Pennica D, Chen J, Yuan J;
 XX WPI, 1999-229533/19.
 DR N-PSDB; AAX52259.
 XX
 PS New isolated human genes and polypeptides used in, e.g. treatment of
 XX Gastrointestinal ulceration.
 XX
 PS Claim 12; Fig 92; 320pp; English.
 XX
 CC AAY13344-403 represent secreted and transmembrane human proteins. The
 CC CDNA sequences are obtained from cDNA libraries, prepared from fetal
 CC lung, fetal kidney, fetal brain, fetal liver and fetal retina. The
 CC encoded polypeptides have specific uses based on their homology to known
 CC polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated
 CC with the preservation and maintenance of gastrointestinal mucosa and the
 CC repair of acute and chronic mucosal lesions (e.g. enterocolitis).
 CC Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
 CC microvillus atrophy), skin diseases associated with abnormal keratinocyte
 CC differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
 CC cell carcinoma of the vulva and gliomas), potent effects on cell growth
 CC and development, diseases related to growth or survival of nerve cells
 CC including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
 CC cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
 CC scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may
 CC be used in the treatment of Usher Syndrome or Atrophla areata; PRO269 can
 CC have therapeutic applications in wound healing and tissue repair; PRO317
 CC blood vessels, or related tissue, e.g. in the heart of genital tract
 CC
 XX Sequence 314 AA:
 QY 14 MGARALLLALLLAAGARKRPSQEAAPLSPGCGRRVYTSIVGGEAELGRMTWQSLR 73
 DB 1 MGARALLLALLLAAGARKRPSQEAAPLSPGCGRRVYTSIVGGEAELGRMTWQSLR 60
 QY 74 LMDSHVCGVSLISRMAITAAHCEFTYSDDLSPSGMWQFQSLSNMSPFSLQAYRYR 133
 DB 61 LMDSHVCGVSLISRMAITAAHCEFTYSDDLSPSGMWQFQSLSNMSPFSLQAYRYR 120
 QY 134 VSNITLSPRYLGNSTPYDIALVKLSAPVTYTHIQIPICQASTFEFENRTDCWVTGWSGK 193
 DB 121 VSNITLSPRYLGNSTPYDIALVKLSAPVTYTHIQIPICQASTFEFENRTDCWVTGWSGK 180
 QY 194 EDEALPSPHTLQEVQVAILINNSMCNHLFLKXSPFRDIFGDMVACGNNQGGDAFCGDSGG 253

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 14:05:54 ; Search time 17 Seconds

(without alignments)
1001.585 Million cell updates/sec

Title: US-09-787-844-2
Perfect score: 1792

Sequence: 1 RROKGRQAGEAMGARGAL.....PSWPLFPPLMLPLGLGPV 327

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1728	96.4	314	1	TEST_HUMAN
2	1147	64.0	324	1	TEST_MOUSE
3	603	33.6	321	1	TRYG_HUMAN
4	580	32.4	343	1	PS88_HUMAN
5	574	32.0	290	1	PR27_HUMAN
6	562.5	31.4	342	1	PS88_MOUSE
7	561.5	31.3	342	1	PS88_MOUSE
8	556.5	31.1	311	1	TRYT_SHEEP
9	544	30.4	273	1	TRYT_SHEEP
10	543	30.3	811	1	TMS6_MOUSE
11	531.5	29.7	317	1	BSS4_HUMAN
12	528.5	29.5	270	1	TRIT_THERON
13	522	29.1	275	1	TRIT_FIG
14	520	29.0	306	1	BSS4_MOUSE
15	515	28.7	275	1	TRB1_HUMAN
16	514.5	28.7	638	1	KAL_MOUSE
17	514	28.7	275	1	TRB2_HUMAN
18	513	28.6	638	1	KAL_HUMAN
19	512	28.6	276	1	MCT6_MOUSE
20	511	28.5	275	1	TRIT_CANFA
21	511	28.5	638	1	KAL_RAT
22	509.5	28.4	273	1	MCT7_RAT
23	509	28.4	811	1	TMS6_HUMAN
24	508.5	28.4	454	1	TMS3_HUMAN
25	507	28.3	273	1	MCT7_MOUSE
26	505	28.2	453	1	TMS3_MOUSE
27	500	27.9	274	1	MCT6_RAT
28	499	27.8	275	1	TRB1_HUMAN
29	490.5	27.4	435	1	TMS4_MOUSE
30	489.5	27.3	418	1	HAT7_HUMAN
31	487.5	27.2	425	1	TMS5_MOUSE
32	483.5	27.0	645	1	PA11_HUMAN
33	478	26.7	437	1	TMS4_HUMAN

34	477.5	26.6	417	1	HBPS_HUMAN
35	472.5	26.4	422	1	DEB1_HUMAN
36	468	26.1	490	1	TMS2_MOUSE
37	466	26.0	436	1	HBPS_MOUSE
38	459.5	25.6	492	1	TMS2_HUMAN
39	457.5	25.5	457	1	TMS5_HUMAN
40	455.5	25.4	415	1	ACRO_PIG
41	455.5	25.4	810	1	PLMN_HUMAN
42	454.5	25.4	1069	1	ENRK_MOUSE
43	454	25.3	338	1	PLMN_HORSE
44	452.5	25.3	436	1	ACRO_MOUSE
45	447	24.9	416	1	HBPS_RAT

ALIGNMENTS

RESULT 1
ID TEST_HUMAN STANDARD; PRT; 314 AA.
AC Q9Y6M0; Q9NS34; Q9P2V6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Testisin precursor (EC 3.4.21.-) (Eosinophil serine protease 1) (ESP-1) (UNQ266/PRO303).
DE PRS21 OR TESTI OR ESP1.
GN Homo sapiens (human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Eosinophil;
RX MEDLINE=99045401; PubMed=9826525;
RA Inoue M., Kanbe N., Kurosawa M., Kido H.;
RT "Cloning and tissue distribution of a novel serine protease esp-1 from human eosinophils.";
RL Biochem. Biophys. Res. Commun. 252:307-312(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=20068805; PubMed=10600542;
RA Inoue M., Ito M., Itoyama T., Kido H.;
RT "Structural analysis of esp-1 gene (PRS 21)." ;
RL Biochem. Biophys. Res. Commun. 266:564-568(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Cervical carcinoma;
RX MEDLINE=9933395; PubMed=1039266;
RA Hooper J.D., Nicol D.L., Dickinson J.L., Eyre H.J., Seaman A.L.,
RA Northey J.F., Stutgen M.A., Douglas M.L., Loveland K.A.,
RA Sutherland G.R., Antalis T.M.;
RT "Testisin, a new human serine proteinase expressed by premeiotic testicular germ cells and lost in testicular germ cell tumors.";
RL Cancer Res. 59:3199-3205(1999).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461760; PubMed=11004480;
RA Hooper J.D., Bowen N., Marshall H., Cullen L.M., Good R., Daniels R.,
RA Stutgen M.A., Northey J.F., Riggs D.R., Kaetner D.L., Ogbourne S.M.,
RA Pera M.F., Jazwinska E.C., Antalis T.M.;
RT "Localization, expression and genomic structure of the gene encoding the human serine protease testisin.";
RL Biochim. Biophys. Acta 1492:63-71(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22887296; PubMed=1295309;
RA Clark H.F., Guney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heidens S.,
RA Huang A., Kim H.S., Kilmowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark W., Robble E., Sanchez C., Schoenfeld J.,
RA Sesnigili S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,

RA	Vandien, J., Watanabe C., Mleand D., Woods K., Xie M.-H., Yansura D., Yi S., Yu G., Yuan T., Zhang M., Zhang Z., Goddard A., Wood W.I., Gadowski P.: "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment." Genome Res. 13:2265-2270(2003).
RL	-1 FUNCTION: Could regulate proteolytic events associated with testicular germ cell maturation.
CC	-1 SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).
CC	-1 ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=3;
CC	Name=1; Synonyms=U/
CC	IsoId=Q9Y6M0-1; Sequence=Displayed;
CC	Name=2; Synonyms=S/
CC	IsoId=Q9Y6M0-2; Sequence=VSP_005389;
CC	Name=3;
CC	IsoId=Q9Y6M0-3; Sequence=VSP_005390;
CC	-1 TISSUE SPECIFICITY: Expressed predominantly in premeiotic testicular germ cells, mostly late pachytene and diplotene spermatocytes.
CC	-1 SIMILARITY: Belongs to peptidase family S1.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
DR	EMBL; AF0583300; AA041588.1; -
DR	EMBL; AB031329; BA883520.1; -
DR	EMBL; AB031330; BA883521.1; -
DR	EMBL; AB031331; BA883522.1; -
DR	EMBL; AF0583301; AAP79019.1; -
DR	EMBL; AF0583301; AAP79020.1; -
DR	EMBL; AY359034; AA089393.1; -
DR	HSSP; P00763; IDPO.
DR	MEROPE; SOL_011; -
DR	Genew; HGNC:9485; PRS821.
DR	MM; 608159; -
DR	GO; GO:0005737; Cytoplasm; TAS.
DR	GO; GO:0005624; Cytoplasmic fraction; TAS.
DR	GO; GO:0005886; Cytoplasm membrane; TAS.
DR	GO; GO:0008236; F-actin-type peptidase activity; TAS.
DR	InterPro; IPR009003; Cys_Ser_trypsin.
DR	InterPro; IPR001254; Peptidase_S1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PRO0722; CHYMOTRYPSIN.
DR	SMART; SMO0020; TRYP_SF1; 1.
DR	PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR	PROSITE; PSS0134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Hydrolase; Serine protease; Glycoprotein; signal; GPI-anchor; Zymogen; Alternative splicing; Lipoprotein.
KW	SIGNAL
FT	PROPEP
FT	CHARIN
FT	PROPEP
FT	ACT_SITE
FT	ACT_SITE
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	LIPID
FT	CARBHYD
FT	CARBHYD

FT	CARBOHYD	273	273	N-LINKED (GLCNAC.) (POTENTIAL).
FT	VARSPLIC	87	88	Missing (in isoform 2).
FT	VARSPLIC	222	235	/FtId=VSP_005389.
FT	SEQUENCE	314 AA; 34884 MW;	E738CF73FB56E98 CRC64;	Missing (in isoform 3).
SQ				/FtId=VSP_005390.
	Query Match	96.4%; Score 1728; DB 1; Length 314;		
	Best Local Similarity	100.0%; Pred. No. 1.2e-147;		
	Matches 314;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	14	MGARGLLALLILARRAGLRKRSSEALPLSGPCGRVITSRIVYGSDADLSNPMWQSRLR	73	
Db	1	MGARGLLALLILARRAGLRKRSSEALPLSGPCGRVITSRIVYGSDADLSNPMWQSRLR	60	
OY	74	LMDSHVCVSLSHRMALTAAHCETASDSDSGMMVPGQLTSMPSIWSQAAYTRYF	133	
Db	61	LMDSHVCVSLSHRMALTAAHCETASDSDSGMMVPGQLTSMPSIWSQAAYTRYF	120	
OY	134	VSNITISPRYLANSFYDALVKLSAPVTYTKHQPICLOASTFEFNRTDCWTGMYTK	193	
Db	121	VSNITISPRYLANSFYDALVKLSAPVTYTKHQPICLOASTFEFNRTDCWTGMYTK	180	
OY	194	EDBALPSPFTIQEVVALINNSMGNHLFLXYSFRKDIFGDMPCAGNAOGGACFCGDSGG	253	
Db	181	EDBALPSPFTIQEVVALINNSMGNHLFLXYSFRKDIFGDMPCAGNAOGGACFCGDSGG	240	
OY	254	PLACKNGKIWIOTGVNSGVCGCPNPFGYTNTISHHFVIQTLMASQMSQPDSPSWLL	313	
Db	241	PLACKNGKIWIOTGVNSGVCGCPNPFGYTNTISHHFVIQTLMASQMSQPDSPSWLL	300	
OY	314	FPELLMALPLIGPV	327	
Db	301	FPELLMALPLIGPV	314	
RESULT 2				
TEST_MOUSE	ID	STANDARD:	PRT,	324 AA.
TEST_MOUSE	SEQHUT; Q9DA14;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Testis in precursor (BC 3.4.21.-) (Trypsin 4).			
GN	PNR821.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/Sv.			
RX	MEDLINE=1133229; PubMed=11231276;			
RA	Scherman A.L., Hooper J.D., Boucaut K.J., Sit M.-L., Webb G.C.,			
RA	Nozmyve U.F., Antalis T.M.,			
RT	"Organization and chromosomal localization of the murine Testis gene			
RT	encoding a serine protease temporally expressed during			
RT	spermatogenesis".			
RL	Eur. J. Biochem. 268:1250-1258(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Testis;			
RX	PubMed=11259427;			
RA	Wong G.W., Li L., Madhusudan M.S., Krilis S.A., Gurish M.F.,			
RA	Richterberg M.E., Sall A., Stevens R.L.,			
RT	"Typease 4, a new member of the chromosome 17 family of mouse serine			
RT	proteases".			
RL	J. Biol. Chem. 276:20648-20658(2001).			
RN	[3]			
RP	SEQUENCE OF 3-324 FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Testis;			
RX	MEDLINE=2108560; PubMed=11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itochi Y.,			